



Antoni Luque Santolaria <aluque@sdsu.edu>

Re: No Subject

Simon Roux <siroux1@gmail.com>
To: "Bas E. Dutilh" <bedutilh@gmail.com>
Cc: Antoni Luque <aluque@sdsu.edu>

Tue, Dec 8, 2020 at 3:13 PM

Hi Bas,

Yikes, this is not great.. A few important information: we have a minimum size cutoff at 1.5kb for all our metagenome-derived virus sequences, so these will not include these really short virus genomes (unfortunately). However what I did in this HQ file was to also include all of Viral RefSeq, assuming (wrongly it seems) that these were real genomes. Looking at the shortest sequences in this file (<250bp), these seem to be all partial sequences but still included in NCBI Viral RefSeq, for whatever reason :-/

Best,
Simon

On Dec 8 2020, at 11:08 am, Bas E. Dutilh <bedutilh@gmail.com> wrote:

Hey Simon,
Toni (in CC) from SDSU is interested in really short virus genomes, so I was checking the shortest 5 sequences in the IMGVR_5.1_HQ_sequences_nucl.fna
They all seem to be fragments, except maybe the 3rd which is a complete 200 bp segment of a segmented virus:
<https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Get&RID=WZK5B11D013>
I thought you mentioned that these already passed CheckV right? Any idea how these slipped through?
Best, Bas

On Sat, Oct 17, 2020 at 9:48 AM Bas E. Dutilh <bedutilh@gmail.com> wrote:

Great thanks a lot Simon! We'll take a look and get back when we have something to show or if we have any questions. Best! Bas

On Sat, Oct 17, 2020 at 1:18 AM Simon Roux <siroux1@gmail.com> wrote:

Hi Bas,

Here are the data we talked about on Wed: <https://drive.google.com/drive/folders/1e7eWEkx5gnSvoYHLG2oXBHBHrJpPJQlg?usp=sharing>
This has the IMG/VR "High-quality" sequences, and the VOG markers we used for taxonomy. There is a README explaining which file is which as well as a few caveats (e.g. about the RefSeq sequences we included). Let me know if you have any questions after going through these (and of course your student is welcome to email me directly with question if any !).

Have a great weekend :-)

Best,
Simon

On Oct 14 2020, at 7:53 am, Bas E. Dutilh <bedutilh@gmail.com> wrote:

Hey Simon,
We could meet through Google: <https://meet.google.com/erw-gvex-osn>
See you in a few minutes!
Bas

On Fri, Oct 9, 2020 at 12:42 PM Bas E. Dutilh <bedutilh@gmail.com> wrote:

Perfect, see you then!
Yes I think DST change is only at the end of the month.
Bas

On Thu, Oct 8, 2020 at 7:57 PM Simon Roux <siroux1@gmail.com> wrote:

Hi Bas,

Thanks ! And always happy to showcase crassphage as a "success story". I feel people relate to it since it's a human gut phage, plus I had no part in any of the papers which I feel makes the message even stronger.

Happy to catch up next week, I'd be available either Wed (Oct 14th) at 8am PST (which I believe will still be 5pm on your side, as daylight saving time is a little later ?).

Best,
Simon

On Oct 8 2020, at 10:19 am, Bas E. Dutilh <bedutilh@gmail.com> wrote:

Hey Simon,
Great talk, thanks for the exposure with the crAssphage part!
Great that you mentioned some of the outstanding challenges in the field!
It might be good to catch up, would you be available somewhere next week? I could make it e.g. Monday or Wednesday at the time you just gave your talk (5 pm our time).
Best! Bas

--

Dr. Bas E. Dutilh
Metagenomics Group
Utrecht Bioinformatics Center

Institute of Biodynamics and Biocomplexity
Theoretical Biology and Bioinformatics, Utrecht University
+31 (0)30 253 4212

--

Dr. Bas E. Dutilh
Metagenomics Group
Utrecht Bioinformatics Center
Institute of Biodynamics and Biocomplexity
Theoretical Biology and Bioinformatics, Utrecht University
+31 (0)30 253 4212

--

Dr. Bas E. Dutilh
Metagenomics Group
Utrecht Bioinformatics Center
Institute of Biodynamics and Biocomplexity
Theoretical Biology and Bioinformatics, Utrecht University
+31 (0)30 253 4212

--

Dr. Bas E. Dutilh
Metagenomics Group
Utrecht Bioinformatics Center
Institute of Biodynamics and Biocomplexity
Theoretical Biology and Bioinformatics, Utrecht University
+31 (0)30 253 4212

--

Dr. Bas E. Dutilh
Metagenomics Group
Utrecht Bioinformatics Center
Institute of Biodynamics and Biocomplexity
Theoretical Biology and Bioinformatics, Utrecht University
+31 (0)30 253 4212